DEVELOPMENT OF RM Y-STR HAPLOTYPE DATABASE FOR ARABIAN PENINSULA POPULATIONS

Rashed Alghafri[1], Mohammad Al Enizi[2], Aqeela Abuidrees[3], Wafaa Al Harrasy[4], Eida Al Mohammed[5], Dina Al Salafi[6], Fatmah Al Tamimi[1], Hamdeh Bin Turkiah[1] Maryam Almheiri[1]
[1]General Department of Forensic Sciences and Criminology, Dubai Police General Head Quarters
[2]Ministry of Interior, General Department of criminal evidences, Identification Department, State of Kuwait
[3]Directorate of Forensic Science Evidence, Public Prosecution, Kingdom of Bahrain
[4] Directorate General of Criminal Inquiries and Investigation, Oman
[5]Ministry of Interior, Doha, State of Qatar
[6]Department of Forensic Science, Sharjah Police General Head Quarters, Sharjah, United Arab Emirates.

Y-chromosomal short tandem repeats (STR) analysis is a valuable tool which has been broadly applied in forensic caseworks especially sexual assaults cases, population studies, genealogical research and kinship analysis. Since the outbreak of rapidly mutating Y-STR (RM Y-STR) loci, these markers were proven to have significantly superior resolution of male lineage differentiation. Therefore, in order to estimate the statistical weight of RM Y-STR haplotype match in the populations of the Arabian Peninsula, a database of detailed haplotypes is required. Forensic laboratories from different Arab's countries collaborated to study and develop an Arabian Peninsula quality controlled RM Y-STR haplotype database. 1228 unrelated, male individuals from six countries; United Arab Emirates, Kingdom of Saudi Arabia, State of Kuwait, State of Qatar, Kingdom of Bahrain and Oman were collected. A subset of these samples was analyzed using Yfiler® allowing the comparison between RM-YSTR and the conventionally used Y-STR markers. Forensic parameter including gene diversity, haplotype diversity, match probability and discrimination capacity were estimated for each population and Arab populations overall. AMOVA analysis was performed between Arab's populations and other populations. Network analysis was also performed using 11 single allelic markers in order to represent the substructure between these populations.